B97-3c: A revised low-cost variant of the B97-D density functional method

Jan Gerit Brandenburg, Christoph Bannwarth, Andreas Hansen, and Stefan Grimme

Citation: The Journal of Chemical Physics **148**, 064104 (2018);
View online: https://doi.org/10.1063/1.5012601
View Table of Contents: http://aip.scitation.org/toc/jcp/148/6
Published by the American Institute of Physics

B97-3c: A revised low-cost variant of the B97-D density functional method

Jan Gerit Brandenburg,[1,2,a)] Christoph Bannwarth,[3] Andreas Hansen,[3] and Stefan Grimme[3]
[1]*Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AH, United Kingdom*
[2]*Thomas Young Centre, University College London, Gower Street, London WC1E 6BT, United Kingdom*
[3]*Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie, Rheinische Friedrich-Wilhelms Universität Bonn, Beringstraße 4, 53115 Bonn, Germany*

(Received 7 November 2017; accepted 21 January 2018; published online 9 February 2018)

A revised version of the well-established B97-D density functional approximation with general applicability for chemical properties of large systems is proposed. Like B97-D, it is based on Becke's power-series ansatz from 1997 and is explicitly parametrized by including the standard D3 semi-classical dispersion correction. The orbitals are expanded in a modified valence triple-zeta Gaussian basis set, which is available for all elements up to Rn. Remaining basis set errors are mostly absorbed in the modified B97 parametrization, while an established atom-pairwise short-range potential is applied to correct for the systematically too long bonds of main group elements which are typical for most semi-local density functionals. The new composite scheme (termed B97-3c) completes the hierarchy of "low-cost" electronic structure methods, which are all mainly free of basis set superposition error and account for most interactions in a physically sound and asymptotically correct manner. B97-3c yields excellent molecular and condensed phase geometries, similar to most hybrid functionals evaluated in a larger basis set expansion. Results on the comprehensive GMTKN55 energy database demonstrate its good performance for main group thermochemistry, kinetics, and non-covalent interactions, when compared to functionals of the same class. This also transfers to metal-organic reactions, which is a major area of applicability for semi-local functionals. B97-3c can be routinely applied to hundreds of atoms on a single processor and we suggest it as a robust computational tool, in particular, for more strongly correlated systems where our previously published "3c" schemes might be problematic. *Published by AIP Publishing.* https://doi.org/10.1063/1.5012601

I. INTRODUCTION

Electronic structure approaches serve as an indispensable complement to experiment by helping us to understand and predict chemical reactivity and the properties of materials as well as guiding the development of new ones. Kohn-Sham density functional theory (DFT)[1,2] is the leading electronic structure technique due to its favorable balance between accuracy and computational cost as well as efficient implementations in modern codes.[3] Modern density functionals improve the short to medium range correlation by the combination of exact constraints with various degrees of parametrization.[4–6] A substantial development in the last decade is the incorporation of dispersion forces in the DFT framework,[7–9] which extends the applicability of DFT to the important class of non-covalently bound molecular complexes and solids. Various approximations make the benchmarking of the proposed methods for the desired application necessary. A few years back, such benchmark studies mainly concentrated on energetics (see, e.g., Refs. 10–14) and those properties are indeed important for chemistry but difficult to compute accurately. Recently, a major focus moved to the description of equilibrium geometries (see, e.g., Refs. 15–17) as those are essential for basically all computational studies. For instance, reliable geometries are needed as a starting point for the investigation of spectroscopic and thermochemical properties[18] or for higher-level wave function theory (WFT) or quantum Monte Carlo (QMC) based studies. Impressive progress has been made in local WFT methods[19–21] and in improved QMC algorithms[22] enabling the description of large molecular complexes[23,24] and organic materials.[25,26] However, it can still be expected that geometry optimizations, vibrational frequency calculations, and molecular dynamics applications will continue to rely on DFT. In this way, DFT should be seen as a complementary method that can be ideally combined with high-level single-point methods. One possible composite approach that combines DFT geometries with local WFT energies has been shown to outperform the W4 standard[27] both in accuracy and in computational efficiency.[28]

For the realistic modeling of supra-molecular or biochemical complexes and organic materials and surfaces, increasingly large systems have to be described at a reasonably accurate electronic structure level. Specifically for large-scale screenings of many structures with the aim of complementing experimental and possibly industrial investigations that set a restricted time frame, a hierarchy of cost-efficient methods is needed. This has been pointed out by recent blind challenges like the statistical assessment of the modeling of proteins and ligands (SAMPL[29,30]) and molecular crystal structure prediction.[31,32] It is crucial for a standard computational tool that it can be applied without the need for a world leading

a)Electronic mail: g.brandenburg@ucl.ac.uk

high-performance computing facility. With this in mind, we have presented a couple of low-cost methods in the past years[15,33,34] and other groups have followed with similar strategies.[35–38] Our current development is triggered by the good performance of the minimal atomic orbital (AO) basis set Hartree-Fock approach with semi-classical corrections for London dispersion and basis set errors (HF-3c).[33] The second composite method in the series is based on a modified Perdew-Burke-Ernzerhof hybrid functional in a modified double-ζ basis set expansion (PBEh-3c).[15] The closely related screened exchange variant (HSE-3c[34]) mainly aims at better numerical robustness for crystals with small band gaps. We sketch a comparison of the different methods according to their basis set size and amount of included Fock exchange in Fig. 1, and an overview of the method ingredients is given in Table I. Central to these developments are well-balanced single particle basis sets that we adjusted carefully for optimum efficiency, specifically regarding good bond lengths for small molecules (see Ref. 39 for an overview). Remaining errors are partially absorbed in the functional parametrization and corrected by classical potentials [D3 (London dispersion),[40] gCP (geometrical counterpoise correction),[41] SRB (short-range basis)[33]]. The functional parametrization determines the semi-local exchange-correlation and has thus major impacts on molecular geometries, thermochemistry, and reactions in general. In contrast, the semi-classical corrections are designed to target one specific aspect, i.e., D3 and gCP impact non-covalent interactions, while the SRB correction mainly effects covalent bond lengths. While PBEh-3c (or its screened exchange variant HSE-3c for solids) is our standard for fast geometry optimizations and interaction energies of organic complexes, HF-3c is preferred for systems where the self-interaction error (SIE) is large. Additionally, it is an order of magnitude faster and analytical (grid free) derivatives make the fast computation of frequencies feasible. These approaches directly profit from efficient integral screening techniques and implementations on modern computer architectures[42–44] enabling full quantum mechanical studies of protein-ligand binding affinities.[45]

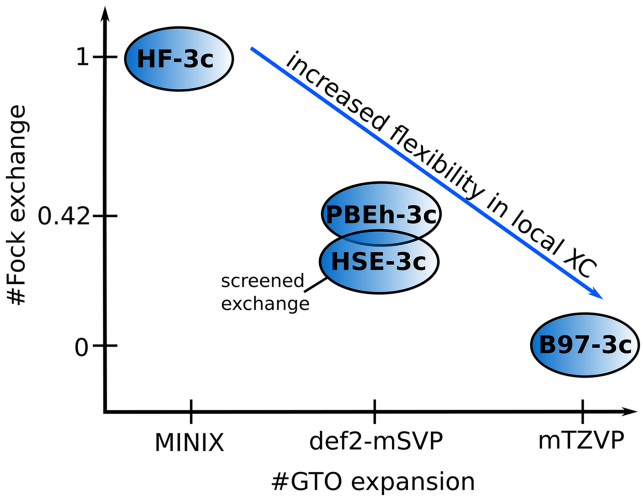

FIG. 1. Sketch of the "3c" composite methods according to their basis set size and amount of included Fock exchange.

TABLE I. Comparison of the hierarchy of "3c" low-cost electronic structure methods.

	HF-3c	PBEh-3c[a]	B97-3c
AO basis set	Minimal	DZ	TZ
No. of parameters in F_x[b]	0	2	4
No. of parameters in F_c[c]	0	1	6
Fock exchange (%)	100	42	0
D3 dispersion	Yes	Yes	Yes
SRB correction	Yes	No	Yes
gCP correction	Yes	Yes	No

[a]HSE-3c identical apart from long-range screening of Fock exchange (range-separation $\omega = 0.11$).
[b]Exchange enhancement compared to local density approximation.
[c]Correlation enhancement, B97-3c has separate parametrization of same-spin and opposite spin.

The main aim here is to complete our hierarchy of "3c" methods by approaching the accuracy of large basis set DFT with a physically sound and numerically well-behaved approach. It is developed along similar lines as PBEh-3c but with the following conceptual and technical changes:

1. Replacing hybrid DFT by a generalized gradient approximation (GGA) for short-range and static electron correlation effects will improve the treatment of electronically more complicated situations like open-shell species or transition metal complexes with partial multi-reference character (see Sec. II A).
2. Replacing the polarized valence-double-zeta basis set by a polarized valence-triple-zeta basis improves the basis significantly, while still being computationally more efficient in most molecular calculations compared to PBEh-3c due to the neglect of Fock exchange (see Sec. II A and Table S1 of the supplementary material).
3. The method is less empirical as only four parameters in the semi-classical correction schemes are freely fitted leading to overall 16 parameters with 13 of them freely optimized (see Sec. II B).

The method is dubbed B97-3c from now on to indicate its origin and the GGA components as usual. The abbreviation highlights the relation to HF-3c and PBEh-3c and more specifically, "3c" stands for the (slightly modified, see below) SRB (for short-ranged basis set errors), the D3 (for dispersion) corrections, and minor modifications of the functional and basis set to ensure consistent bond lengths across the periodic table as well as modifications of the semi-local exchange and correlation parametrization. B97-3c will be particularly useful for systems with partial multi-reference character, where both HF-3c and PBEh-3c are problematic due to the large amount of Fock exchange. As the basis set is increased to triple-ζ quality, we can profit from a more flexible exchange correlation functional and the B97 form was chosen based on the success of the B97-D functional[46] and the good performance of newer variants like the B97M-V.[47]

We start with a full definition of the new B97-3c method in Sec. II followed by computational and implementation details in Sec. III. B97-3c is widely tested and compared to the other "3c" methods as well as established density functionals in Sec. IV, where we benchmark molecular geometries

(Sec. IV A), non-covalent interactions (Sec. IV B), main group thermochemistry and kinetics (Sec. IV C), metal-organic reactions (Sec. IV D), and lattice energies and geometries of molecular solids (Sec. IV E). We give a final recommendation of the applicability of B97-3c in our conclusions (Sec. V).

II. THEORY

A. B97-3c construction

1. B97 Taylor expansion

The starting point is a standard GGA density functional with Taylor expansions of the semi-local exchange and correlation as introduced by Becke in 1997.[48] We briefly review the method and our extensions to treat large systems efficiently including attempts to avoid extensive basis set superposition error (BSSE). The B97 functional is based on a remapping of the reduced gradient variable

$$s_\sigma = \frac{\nabla n_\sigma}{n_\sigma^{4/3}}, \tag{1}$$

where n is the electron density and σ denotes α or β spin. The density dependent part of the exchange-correlation functional is given as

$$E_{XC}^{\mathrm{B97}} = E_X + E_{C\alpha\beta} + \sum_\sigma E_{C\sigma\sigma}, \tag{2}$$

where X and C denote exchange and correlation contributions, respectively, given by

$$E_X = \sum_\sigma \int e(n_\sigma) g_{X\sigma}(s_\sigma^2) dr, \tag{3}$$

$$E_{C\alpha\beta} = \int e(n_\alpha, n_\beta) g_{C\alpha\beta}(s_{av}^2) dr, \tag{4}$$

$$E_{C\sigma\sigma} = \int e(n_\sigma) g_{C\sigma\sigma}(s_\sigma^2) dr, \tag{5}$$

where $e(n)$ in Eqs. (3)–(5) are local energy densities of a uniform electron gas, g denote gradient correction factors, and $s_{av}^2 = 1/2(s_\alpha^2 + s_\beta^2)$. The correction factors are expanded in a power series in the remapped variable $u(s^2)$ (spin-subscripts omitted)

$$g(s^2) = \sum_{j=0}^{k} c_j u^j(s^2). \tag{6}$$

Already from Becke's work[48,49] and from extensive work of Head-Gordon and co-workers, it is known that $k = 2$ (three terms) is a good compromise between flexibility and "robustness" for a GGA functional and that $k = 4$ as in previous B97 parametrizations[50–52] seems to represent some kind of over-fitting. For the three different parts in E_{XC}, the following forms for g are used:

$$u_{X\sigma}(s_\sigma^2) = \frac{\gamma_{X\sigma} s_\sigma^2}{1 + \gamma_{X\sigma} s_\sigma^2}, \tag{7}$$

$$u_{C\alpha\beta}(s_{av}^2) = \frac{\gamma_{C\alpha\beta} s_{av}^2}{1 + \gamma_{C\alpha\beta} s_{av}^2}, \tag{8}$$

$$u_{C\sigma\sigma}(s_\sigma^2) = \frac{\gamma_{C\sigma\sigma} s_\sigma^2}{1 + \gamma_{C\sigma\sigma} s_\sigma^2}. \tag{9}$$

As in B97-D, the linear parameters c are re-determined here by a least-squares fit procedure, while the non-linear parameters γ are taken from Becke's work.[48]

2. Basis set modifications

The density functional B97 is evaluated in a medium-sized Gaussian orbital basis set of triple-ζ quality, dubbed mTZVP in the following. The Ahlrichs basis set def-TZVP[53] was chosen as a starting point with a few individual modifications for the elements H–Ar (see Table II and Table S1 of the supplementary material). Notably is the reduced polarization on hydrogen atoms for significant speed up and additional polarization functions on oxygen atoms, which improve the description of strong hydrogen bonds. For the remaining heavier elements, the standard def2-TZVP basis set is used with corresponding matching effective core potentials of the Stuttgart-Cologne type as implemented originally in TURBOMOLE. A full definition is given in the supplementary material.

3. Semi-classical correction potentials

The total B97-3c Kohn-Sham energy is calculated as

$$E_{\mathrm{tot}}^{\mathrm{B97\text{-}3c}} = E_{\mathrm{tot}}^{\mathrm{mB97}} + E_{\mathrm{D3}} + E_{\mathrm{SRB}}, \tag{10}$$

where mB97 denotes the re-fitted version of E_{XC}^{B97} (in the special triple-zeta basis set mTZVP) and the first correction term E_{D3} is the London dispersion energy from the D3 scheme[7,40]

$$E_{\mathrm{disp}}^{(\mathrm{D3})} = -\frac{1}{2} \sum_{n=6,8} \sum_{A,B}^{\mathrm{atoms}} s_n \frac{C_n^{AB}}{R_{AB}^n} \cdot f_n^d(R_{AB}) - \frac{1}{6} \sum_{A,B,C}^{\mathrm{atoms}} \frac{C_9^{ABC}}{R_{ABC}^9} \cdot f_9^d(R_{ABC}, \theta_{ABC}), \tag{11}$$

TABLE II. Contraction scheme of the revised basis set mTZVP for the elements H, C, N, and O compared to the standard def-TZVP basis set. Contractions for all elements H–Rn and explicit exponents are given in Table S1 of the supplementary material.

	def-TZVP	mTZVP
Element	Contraction	
H	(5s1p)/[3s1p] {311/1}	(5s)/[3s] {311}[a]
C	(11s6p1d)/[5s3p1d] {62 111/411/1}	(11s6p1d)/[5s3p1d] {62 111/411/1}[b]
N	(11s6p1d)/[5s3p1d] {62 111/411/1}	(11s6p1d)/[5s3p1d] {62 111/411/1}[b]
O	(11s6p1d)/[5s3p1d] {62 111/411/1}	(11s6p2d)/[5s3p2d] {62 111/411/11}[c]

[a]Polarization function removed for efficiency reasons.
[b]Identical to original TZVP basis.
[c]Second polarization function added for improved hydrogen bonding.

with atomic distances R_{AB}, geometrically averaged distance R_{ABC}, angles of the atomic triangles θ_{ABC}, dispersion coefficients C_n, and damping functions f_n^d. In D3 the dynamic polarizabilities of molecular hydrides are calculated via time-dependent DFT for reference systems, and a modified Casimir-Polder integration yields the atom-pair C_6^{AB} value. Higher-order dipole-quadrupole pair-terms and Axilrod-Teller-Muto (ATM) type[54,55] three-body terms are calculated via recursion relations and averages, respectively, from the corresponding C_6 coefficients. The importance of many-body dispersion interactions has been recently analyzed by various groups.[56–58] The two-body contributions C_6 and C_8 can be scaled by parameters s_6 and s_8. A short-range damping function switches from the semi-local correlation to the correction scheme and contains two parameters (a_1 and a_2) for the rational Becke-Johnson damping[59–61] and one parameter in the zero-damping variant (r_{s6}). The second correction term E_{SRB} denotes a slight modification of the short-range basis (SRB) set correction developed for HF-3c,[33]

$$E_{\mathrm{SRB}} = -\frac{q_{\mathrm{scal}}}{2} \sum_{A,B}^{\mathrm{atoms}} \sqrt{Z_A Z_B}\, \mathrm{e}^{-r_{\mathrm{scal}} \frac{R_{AB}}{R_0^{AB}}}, \tag{12}$$

where q_{scal} and r_{scal} are the fitting parameters, R_{AB} is the distance of the atom-pair AB, $Z_{A/B}$ are their atomic numbers, and R_0^{AB} are the D3 damping radii.[40] The SRB correction mainly influences covalent bond lengths that are systematically too long for most (meta-)GGA density functionals, with the exception of M06L and SCAN as shown in Ref. 62. In contrast to the previously developed methods, we do not use an explicit counterpoise correction potential (e.g., gCP). It has been included in an initial fit, but no significant improvement could be achieved. Thus, the potential has been removed and remaining basis set errors are absorbed by the functional in the least-square minimization process (see below). As shown in Sec. IV B, no overbinding tendency for non-covalently bound systems of varying size can be seen, which demonstrates the validity of our approach.

B. Parameter fitting

The expansion parameters [Eq. (6)] of the B97-3c functional have been determined by a damped least-squares optimization (Levenberg-Marquardt algorithm[63,64]) with respect to energetic deviations of a training set of systems. The D3 dispersion correction was included in this step. The set of molecules and reactions consisted of 20 atomization energies, 43 chemical reactions, and 27 non-covalently bound complexes. The systems and corresponding reference data have been taken mainly from subsets of the general main-group, thermochemistry, kinetics, and non-covalent interactions (GMTKN55) database;[12,65] details on the set are given in the supplementary material. Four total atomic energies were added to this set, which resulted in a more stable least-square parameter optimization. Geometries of 17 small molecules are included by considering the total gradient on reference structures, mainly influencing the parameter of the SRB correction. The starting parameters for E_X and E_C have been taken from the well-established B97-D functional[46] and they provided good results across all chemical properties and a recent global search confirmed them to be close to the global minimum.[47] The final B97-3c parameters including the D3 damping parameters are given in Table III. The values of the electronic part are similar to those used in B97-D with deviations of 7% with the exception of the zeroth order same-spin correlation ($c_{\sigma\sigma}^0$) that is about doubled. The increase in the latter apparently makes the functional more repulsive, which effectively corrects for the residual BSSE. The final B97-3c RMSD on the fit set (excluding geometries) is 38.5 kcal/mol. This has to be interpreted carefully, as some data points like the total energies give a large energetic contribution but only have a small impact on the method performance for chemical properties. Benchmark data focusing on actual chemical problems will be provided in Sec. IV.

TABLE III. Empirical parameters of the B97-3c composite electronic structure method.

	Parameter			
Contribution	γ	c_x^0	c_x^1	c_x^2
E_X	0.004[a]	1.076 616	−0.469 912	3.322 442
		$c_{\sigma\sigma}^0$	$c_{\sigma\sigma}^1$	$c_{\sigma\sigma}^2$
E_C^{SS}		0.543 788	−1.444 420	1.637 436
		$c_{\alpha\beta}^0$	$c_{\alpha\beta}^1$	$c_{\alpha\beta}^2$
E_C^{OS}		0.635 047	5.532 103	−15.301 575
	s_6	s_8	a_1[b]	a_2[b]
E_{D3}	1.00[a,c]	1.50[a,d]	0.37	4.10
			q_{scal}	r_{scal}
E_{SRB}			10.00	0.08

[a]Constrained value.
[b]Zero-damping variant with parameter $r_{s6} = 1.06$.
[c]Set to unity to ensure the correct long-range limit.
[d]Rounded value taken from PBE38.[40]

The implementation of B97-3c into any quantum chemistry code is straightforward if the software infrastructure for a GGA functional is available. For the D3 and SRB corrections, freely available codes can be downloaded from our website.[66]

III. TECHNICAL DETAILS

A. Implementation and computational settings

The B97-3c composite electronic structure method has been implemented in ORCA 4.0.1[67] and TURBOMOLE 7.2.[68,69] Additional implementations are available in a development version of CRYSTAL17[70] and CP2K 2.7.[71–73] Most molecular calculations have been conducted with ORCA, while CRYSTAL has been mainly applied to periodic systems. TURBOMOLE has been used as cross-check for both molecular and periodic cases. When applicable, the RI-J approximation was used[74–76] with appropriate auxiliary basis sets.[77] For the semi-local exchange-correlation part, the medium-sized numerical quadrature grids are used (grid5 in ORCA, grid m4 in TURBOMOLE, LGRID in CRYSTAL)[78] and tight self-consistent field (SCF) convergence criteria were applied in geometry optimizations. In ORCA, the method is invoked with the compound keyword "b97-3c," while in TURBOMOLE and CRYSTAL, it is used as a functional name and the appropriate basis set mTZVP has to be chosen manually. Except

TABLE IV. Comparison of timings for the electronic structure methods considered in this study. We give the number of involved (spherical) atomic orbitals, and the computation times (hh:mm:ss) refer to the time needed for a single-point calculation on a single core (Intel(R) Xeon(R) CPU E3-1270 v5 @ 3.60 GHz, 64 GB RAM) using the ORCA program (version 4.0.1). The considered systems are phenanthrene ($C_{14}H_{10}$), [7]helicene ($C_{30}H_{18}$), and [16]helicene ($C_{66}H_{36}$).

	Phenanthrene		[7]helicene		[16]helicene	
	No. of orbitals	Time	No. of orbitals	Time	No. of orbitals	Time
HF-3c	80	00:00:02	168	00:00:16	366	00:05:02
B97-3c[a,b]	296	00:00:45	624	00:03:18	1362	00:18:26
PBEh-3c[a,b]	230	00:02:05	486	00:13:29	1062	02:04:41
BP86-D3^{atm}/def2-TZVP[a,b]	494	00:01:27	1038	00:07:09	2262	00:42:56
PBE-D3^{atm}/def2-QZVP[a,b]	1098	00:04:11	2250	00:21:10	4842	02:30:44
TPSS-D3^{atm}/def2-QZVP[a,b]	1098	00:06:20	2250	00:30:12	4842	02:50:37
HF/def2-QZVP	1098	04:09:36	2250	36:41:06	4842	401:50:54

[a]The numerical integration grid *grid4* has been used.
[b]The RI-J approximation has been used.

for HF-3c, the "3c" methods use the D3 correction with the Axilrod-Teller-Muto type three-body term, while otherwise it is indicated by the superscript atm. In some cases (e.g., for anions), it is necessary to add diffuse functions for specific atoms as usual. According to some test calculations for B97-3c, it seems to be the most robust strategy to keep the standard SRB and D3 parameters in this case. For the computation of vibrational (free) energies, we propose to use B97-3c frequencies unscaled as usual for GGAs.

B. Computational efficiency

In Table IV, the computation times required for a single-point calculation with different methods are given. For this purpose, phenanthrene, [7]helicene, and [16]helicene are used. The number of contracted basis functions in the MINIX basis set, which is employed in HF-3c, is smaller by about one order of magnitude compared to the large def2-QZVP basis. This leads to tremendous computational savings of about 3–4 orders of magnitude. This makes HF-3c the fastest of all methods considered. However, the less beneficial scaling (formally $\propto N^4$) with increasing system size becomes apparent from the relative speed-up obtained for larger systems compared to the pure density functional methods, which formally scale with N^3 due to the RI-J approximation. The proposed B97-3c method, which employs a much larger basis set, ranks second. For the largest system considered here ([16]helicene), it is only 3–4 times slower compared to HF-3c. It is also faster by a factor of six for this system than the PBEh-3c composite procedure, and thus, it fills the gap between our existing "3c" methods in terms of the computational cost.

The applied modifications in the mTZVP basis set furthermore make the B97-3c method for an organic molecule faster by a factor of two to three compared to a combination of a GGA with the standard def2-TZVP basis (cf. BP86-D3^{atm}/def2-TZVP). The specific modifications of the basis set and the functional in B97-3c result in rather accurate energetic properties (see GMTKN55 data discussion). Thus, the smaller computational demands compared to the similarly performing (meta-)GGA-D3^{atm}/def2-QZVP combinations on the GMTKN55 set also enable fast computation of reasonable single-point energies, whenever the use of a semi-local functional is appropriate.

IV. RESULTS AND DISCUSSION

The proposed composite method shall be applicable to all kinds of organic, inorganic, and organo-metallic systems. Due to its low-cost character, the main focus is on structural properties and non-covalent interaction energies in large systems. At the same time, thermochemical properties shall be within good to very good accuracy as tested with the huge general main-group, thermochemistry, kinetics, and non-covalent interactions (GMTKN55) database.[12,65] We will focus on the "3c" type methods HF-3c, PBEh-3c, and B97-3c. To highlight the impact of the B97 parametrization and the SRB scheme, we additionally show the original B97-D3^{atm} in the mTZVP basis set and the new functional (termed mB97-D3^{atm}) without the SRB correction. Results from the standard functionals PBE,[79] TPSS,[80] and BP86[81,82] (all corrected for London dispersion) as well as plain HF and MP2 are given in the supplementary material.

A. Geometries

1. Bond length

As we used part of the light main group bond set (LMGB35) in the training of the SRB correction, we start the geometry evaluation with the independent CCse21 set of Barone and co-workers.[17] It consists of semi-experimental equilibrium structures of small semi-rigid organic molecules. The reference geometries are derived by a least-squares fit of the structural parameters to the experimental ground-state rotational constants corrected by vibrational contributions computed at the coupled cluster level (CCSD(T) with at least triple-ζ basis sets). Thus, the set is ideally suited to test the predictive quality of different electronic structure methods for bond lengths, which has been previously done for a wide set of traditional density functionals.[83] Results from our composite methods are shown in Fig. 2 and statistical measures are given in Table V. HF-3c performs similar to plain HF or BLYP (both evaluated in large basis sets) highlighting the systematic error

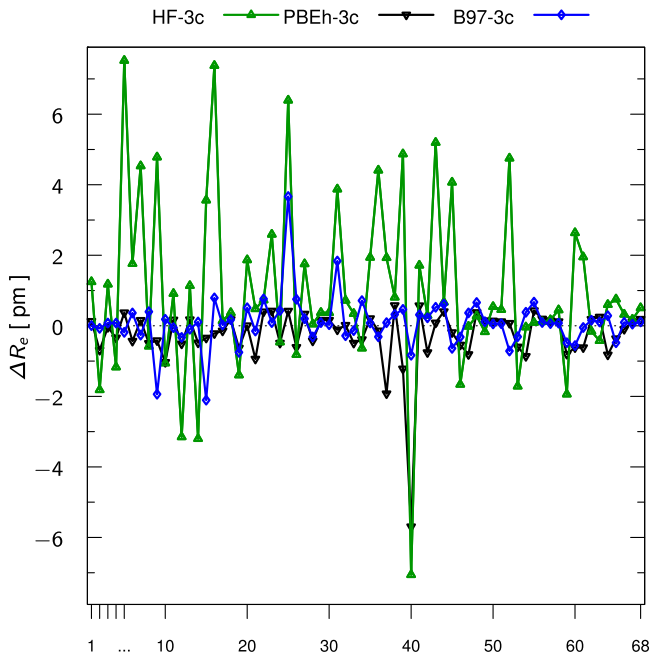

FIG. 2. Deviations of bond distances in the 21 molecules (68 bonds) of the CCse21 set.[17] The numbering of all bonds and individual results of additional functionals are given in the supplementary material.

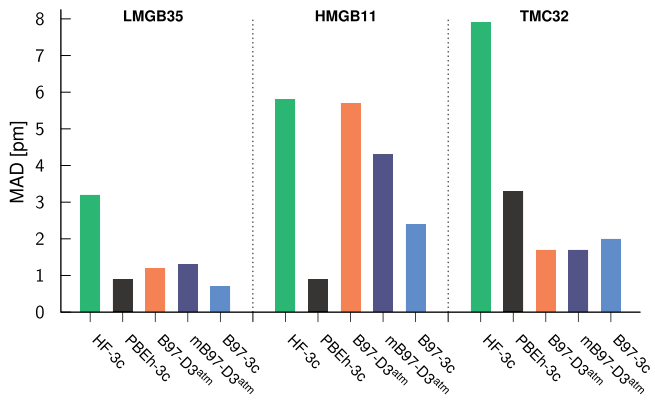

FIG. 3. Mean absolute deviations of various methods for different bond distances separated into light main group bonds (LMGB35), heavy main group bonds (HMGB11), and transition metal complexes (TMC32).

compensation between too short bonds from HF and elongated bonds from the minimal basis set.[84] A substantial improvement is made with PBEh-3c, which has been shown in Ref. 15, and is confirmed on the CCse21 set with small mean absolute deviation (MAD) of the considered bond lengths below 0.5 pm. Typical non-hybrid functionals have bond length MADs around 1.0 pm, and our revised mB97-D3 functional in the reduced mTZVP basis set performs similar. Including the SRB correction yields very accurate bond lengths with an MAD of only 0.4 pm, outperforming all previously tested non-hybrid functionals and is even competitive to the best global and range-separated hybrid functionals (PBE0 and ωB97X).[83,85] The SRB correction slightly deteriorates the agreement for the bond angles by 0.03°. The analysis of bond angles mainly indicates an increased accuracy when the basis set is enlarged and the averaged structure root mean square deviations (RMSDs) of 0.7 pm demonstrate the reliable geometries of both PBEh-3c and B97-3c.

As the new composite method should be widely applicable to all kinds of chemical structures, we additionally test established sets of light main group bonds (LMGB35[15]), heavy main group bonds (HMGB11[15]), and 3d-transition metal complexes (TMC32[86]). MADs of the composite schemes are shown in Fig. 3. The global hybrid PBEh-3c yields excellent heavy main group bond lengths but is slightly worse for transition metal bonds compared to non-hybrid functionals. This is mainly due to the large amount of Fock exchange of 42% and we thus expect the best methods found previously for the CCse21 set[83] that rely on large portions of Fock exchange and second-order perturbation theory (e.g., the double hybrid B2PLYP[87]) to encounter the similar problems. B97-3c has an overall very balanced performance with MADs of 0.7, 2.4, and 2.0 pm for LMGB35, HMGB11, and TMC32, respectively. While transition metal bonds are hardly affected by the SRB correction, it yields a substantial improvement for the bonds of main group elements.

2. Molecular structures

In order to investigate molecular structures of medium-sized molecules, we use the rotational constants from the ROT34 set.[88,89] Rotational constants can be measured accurately in the gas phase at low temperatures and are directly related to the inverse size (and shape) of a molecule, $B \propto r^{-2}$. Zero-point vibrational effects have been removed for convenient benchmarking of electronic structure methods. We show the statistics on the ROT34 set converted to normal error distributions in Fig. 4. While HF-3c yields similar results to HF in a large basis set expansion, both PBEh-3c and B97-3c perform substantially better. Notably is the significant effect of the SRB scheme, correcting for the systematically too long bonds

TABLE V. Statistical deviations of bond lengths and bond angles in the 21 molecules (68 bonds, 42 angles) of the CCse21 set.[17]

	Δ(bond length) (pm)				Δ(angle) (deg)				Structure
	MD	MAD	SD	AMAX	MD	MAD	SD	AMAX	RMSD (pm)
HF-3c	1.0	1.8	2.5	7.5	0.2	0.9	1.2	3.4	2.27
PBEh-3c	–0.3	0.5	0.8	5.7	0.0	0.6	0.8	1.9	0.70
B97-D3^{atm} [a]	0.9	0.9	0.8	5.0	0.0	0.6	0.7	1.3	1.16
mB97-D3^{atm} [a]	0.9	0.9	0.8	4.3	0.0	0.5	0.6	1.3	1.16
B97-3c[a]	0.1	0.4	0.7	3.7	0.0	0.6	0.7	1.4	0.71

[a]All evaluated in the same mTZVP basis set.

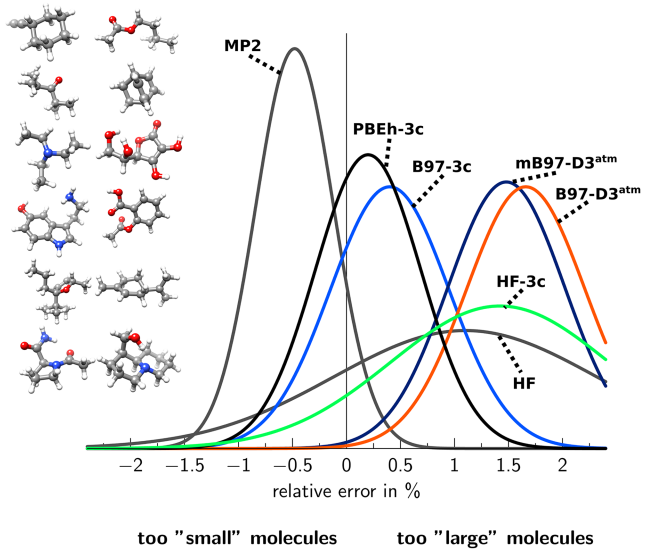

FIG. 4. Normal distribution of the relative errors in the computed rotational constants B_e for the ROT34 benchmark set with various theoretical methods. HF and MP2 results are taken from Ref. 88. The inset shows the molecules of this set.

of the B97 GGA. The molecular structures are still about 0.4% too large, but this effect is negligible compared to nearly all other non-hybrid functionals (with the exception of SCAN-D3 and M06L[90]) and smaller than the systematically too compact MP2 structures.

B. Non-covalent interactions

Specifically aiming at larger systems, the non-covalent interactions become increasingly important. It has been shown previously that intrinsically more repulsive semi-local density functionals can be better combined with corrections for London dispersion.[91,92] Here, we show that this also holds for our new B97-3c composite method.

The S66x8 set[93] is the established standard in testing non-covalent interactions with new coupled cluster with singles, doubles, and perturbative triples [CCSD(T)] reference values from Ref. 94. The set consists of the equilibrium geometry and seven additional varied center-of-mass distances. We use cubic splines to interpolate the 66 potential curves and extract the equilibrium center-of-mass distance and interaction energy. Even with the reduced basis sets, we get the typical dispersion corrected DFT accuracy with MADs for the interaction energies below 0.5 kcal/mol and B97-3c performs best with an MAD of 0.33 kcal/mol. The SRB correction only slightly influences the non-covalent interactions, which is important for the shown good performance. Thus, it can be readily combined with the semi-classical D3 correction, which is less trivial for the over-attractive metaGGAs like SCAN and M06L. The non-covalent distances as measured with the equilibrium center-of-mass distance are all reasonably good with typical errors of 5-6 pm. The HF-3c distances are systematically too short, while all other methods yield slightly too large distances. Again, we tested the impact of the basis set expansion. The MAD on the S66 energies increases to 0.4 and 1.5 kcal/mol for B97-3c in a quadruple and double-zeta basis set, respectively. A good performance on the S66 set is certainly

TABLE VI. Statistical deviations of interaction energies ΔE_{int} and center-of-mass distances ΔR_{com} from cubic spline interpolation of the 66 molecular dimers (S66x8[93,94]). Interactions in large complexes are tested by the S30L[95] and L7[96] benchmark sets.

	S66 ΔE_{int} (kcal/mol)				S66 ΔR_{com} (pm)			
	MD	MAD	SD	AMAX	MD	MAD	SD	AMAX
HF-3c	0.0	0.4	0.5	1.9	−5.0	5.3	5.0	21.9
PBEh-3c	−0.1	0.5	0.6	2.1	3.3	5.9	7.0	16.7
B97-D3$^{\text{atm}}$ [a]	−0.3	0.4	0.4	1.5	4.3	5.1	5.4	20.3
mB97-D3$^{\text{atm}}$ [a]	0.0	0.3	0.4	1.4	6.5	6.7	6.9	28.4
B97-3c[a]	0.0	0.3	0.5	1.4	6.1	6.5	7.0	28.4
	S30L ΔE_{int} (kcal/mol)				**L7** ΔE_{int} (kcal/mol)			
HF-3c	−3.1	5.2	7.4	25.1	0.9	1.2	1.4	3.6
PBEh-3c	0.0	2.6	3.5	10.5	1.3	1.8	2.2	4.5
B97-D3$^{\text{atm}}$ [a]	−5.0	6.6	4.3	15.4	−2.3	2.3	0.8	3.5
mB97-D3$^{\text{atm}}$ [a]	−1.6	3.2	4.0	10.5	0.2	0.8	0.9	1.4
B97-3c[a]	−1.8	3.2	4.0	10.4	0.1	0.8	0.9	1.3

[a] All evaluated in the same mTZVP basis set.

a requirement for a good electronic structure method, but further cross-checks on larger systems are crucial. We choose a set of large host-guest complexes (S30L[95,97]) and the large molecular complexes from the L7 set.[96] We use new reference values (see the supplementary material, Tables S20–S22) based on localized CCSD(T) at the estimated basis set limit [DLPNO-CCSD(T)[23] in its sparse matrix implementation[19] employing the CBS* protocol as described in Ref. 98]. The statistical performance on all three sets is given in Table VI. The good result for B97-3c is confirmed by the low MADs of 3.2 and 0.8 kcal/mol for S30L and L7, respectively. For the L7 set, this is better than any other density functional tested so far, e.g., PBE-D3$^{\text{atm}}$ and TPSS-D3$^{\text{atm}}$ have MADs of 2.6 and 1.1 kcal/mol.[62] For the larger systems, the systematic overbinding of the B97-D3$^{\text{atm}}$ functional in the mTZVP basis set indicates a remaining BSSE. This is effectively absorbed in the functional re-parametrization for the B97-3c and all mean deviations are comparably small. While comparing the deviations on the three NCI benchmark sets S66, S30L, and L7, one has to consider the difference in mean binding energies of 5.5, 37.5, and 16.7 kcal/mol. A more balanced comparison is given by the mean relative deviations shown in Fig. 5. The new composite methods yield very good interaction energies with errors well below 11% for all sets. Again we note that counterpoise corrections are not necessary for B97-3c, which makes the methods readily applicable and a factor of two faster than related methods needing an explicit counterpoise correction.

C. Main group thermochemistry and kinetics

The performance of B97-3c for molecular structures and non-covalent interactions has been investigated in Secs. IV A and IV B. Recently, the GMTKN55 benchmark database[65] has been compiled.

The weighted total MADs (WTMAD) for this database and the individual subset results obtained with five different methods are plotted in Fig. 6 (the second weighting scheme presented in Ref. 65 is used, denoted WTMAD-2). We excluded HF-3c from this figure since the WTMAD-2

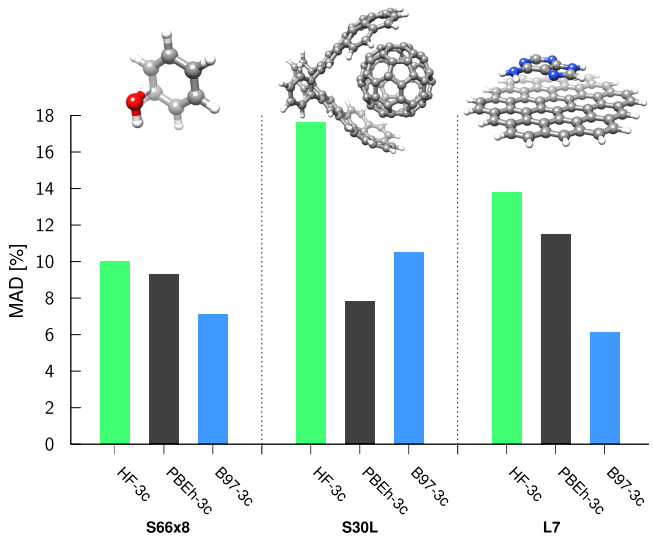

FIG. 5. Mean absolute relative deviations (MADs) of the three composite methods for the noncovalent interaction benchmark sets S66x8[93] (molecular dimers), S30L[95] (host-guest complexes), and L7[96] (large molecular complexes). The insets show one typical system of each set; a full list is given in the supplementary material.

values are significantly larger ranging from about 23 to 44 kcal/mol, precluding a reasonable depiction with the other methods (see the supplementary material for the complete figure). It should be kept in mind that HF-3c, being devoid of short- and medium-range correlation effects, does perform rather badly for basic properties and barriers [labeled as (a) and

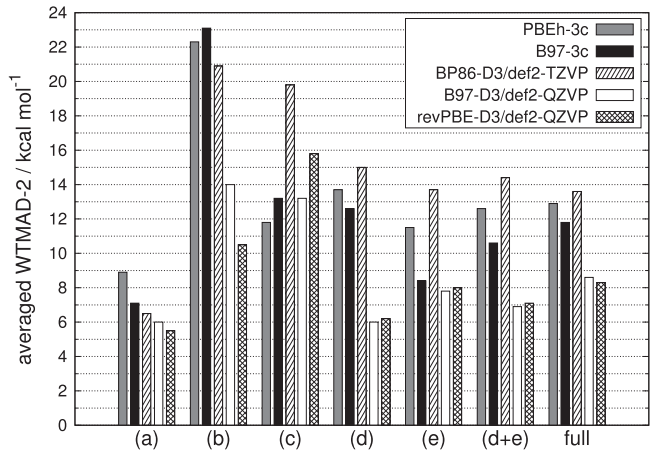

FIG. 6. Averaged weighted mean absolute deviations (WTMAD-2, see Ref. 65) in kcal/mol for the PBEh-3c and B97-3c methods on the GMTKN55 benchmark database. Due to the significantly larger WTMAD-2 values (by a factor of 2–4), HF-3c is neglected in this plot (see the supplementary material). For comparison, the results for the B97-D3 functional and the best-performing GGA on the GMTKN55 (revPBE-D3) with the large def2-QZVP (additional diffuse s- and p-functions for non-hydrogens in WATER27, G21EA, AHB21, and IL16 sets[65]) are also given. As a competitor to B97-3c, results with BP86-D3 with the larger def2-TZVP basis set are also shown. All methods employ the pair-wise D3 dispersion corrections with Becke-Johnson damping. The latter is an inherent component of the 3c approaches, and PBEh-3c and B97-3c are also augmented with the D3 ATM term. The values are displayed for the various categories of the GMTKN55 set: (a) basic properties and reactions of small systems, (b) isomerisations and reactions of large systems, (c) barrier heights, (d) intermolecular non-covalent interactions, and (e) intramolecular non-covalent interactions. The combined values for all noncovalent interactions (d + e), as well as for the full GMTKN55 database, are also shown.

(b) in Fig. 6], while the description of non-covalent interactions is more reasonable, which has been discussed in Sec. IV B.

The WTMAD-2 obtained with PBEh-3c and B97-3c is shown along with the results for the competing GGA/triple-ζ scheme BP86-D3/def2-TZVP. For further comparison, the B97-D3 method and the best-performing GGA on the full GMTKN55 database, revPBE-D3, are shown, which were employed with the large def2-QZVP basis sets (partially augmented with diffuse functions[65]).

As expected from the large GMTKN55 benchmark study,[65] the hybrid functional-based PBEh-3c method performs significantly better for barriers [(c) in Fig. 6] compared to the GGA-based methods suffering from the self-interaction error. revPBE-D3/def2-QZVP and especially BP86-D3/def2-TZVP perform significantly worse than PBEh-3c by about 4 and 7 kcal/mol, respectively. However, the functional form of B97-D seems to work comparably well for barriers, as the WTMAD-2 of B97-3c and B97-D3/def2-QZVP is only slightly larger by about 1.5 kcal/mol compared to PBEh-3c.

For basic properties and reactions of small systems [(a) in Fig. 6], PBEh-3c gives the largest WTMAD-2 of the methods considered, which is likely a result from the use of a too small basis set. For the other methods, it is also observable that the method, which is combined with a larger, more polarized basis set, yields a smaller WTMAD-2. The included quantities in this subset include atomization energies and covalent binding energies, which are described much better if a large and polarized basis set is employed. For example, the performance of the B97-D functional is also only slightly worse (0.5 kcal/mol) compared to that of revPBE, while B97-3c shows a WTMAD-2 which is larger by 1.5 kcal/mol. Given the similar functional form of B97-D and B97-3c, the differences in the one-particle basis set seem to be mainly responsible for the observed differences. This is also observable for the other sets that include bond cleavage/formation reactions, i.e., (b) and (e) in Fig. 6. For the latter intramolecular noncovalent interaction sets, however, B97-3c performs exceptionally well and hardly any difference from the computationally more demanding B97-D3/def2-QZVP approach can be seen.

For the entire GMTKN55 set with 1500+ reactions, the performance of B97-3c is the best of the herein considered "low-cost" approaches. Its WTMAD-2 is smaller by about 2 kcal/mol compared to BP86-D3/def2-TZVP. As discussed above, the discrepancy of roughly 3 kcal/mol with respect to B97-D3/def2-QZVP results from the less complete basis set (CBS), which, however, is important in the description of bond formation/cleavage reactions that contribute significantly to the GMTKN55 but results also in significantly increased computation times. None of the GGA functionals can deliver very accurate thermochemistry and reaction barriers. For higher accuracy, hybrid and double hybrid functionals can be employed[65] or many-body methods like CCSD(T) or QMC, if affordable.

D. Metal-organic reactions

The GMTKN55 benchmark database includes only main group elements and hence the performance of B97-3c for

transition metal chemistry has to be assessed otherwise. We have already shown in Sec. IV A that B97-3c is very promising for the geometries of 3d-transition metal complexes. Here, we test if this also holds for metal organic thermochemistry. The problems of generating reliable reference energies are well known for cases with strong multi-reference character.[99] Still, non-hybrid functionals have been shown to be rather robust for many 3d transition metal reactions;[100] see Ref. 101 for a comprehensive review on the applicability of DFT for transition metal chemistry. Here, we employ a recently compiled benchmark set with 41 typical metal organic reactions (MORs).[102] It consists of reactions often found in catalytic cycles with the largest molecules containing about 120 atoms and reactions involving the transition metals Ti, Cr, Mn, Fe, Co, Ni, Mo, Ru, Rh, Pd, W, Ir, and Pt. The set includes σ- and π-donor complexation, oxidative addition, and ligand exchange reactions with reaction energies varying from 27 kcal/mol to −66 kcal/mol (smallest value of −1.9 kcal/mol, the average reaction energy is about −30 kcal/mol). All systems are closed-shell and neutral without significant multi-reference character. The reference values are computed at the DLPNO-CCSD(T) level[23] in its sparse maps implementation[19] as available in ORCA 4.0.0[67] using the TightPNO[103] settings and the TightSCF convergence criteria for the HF energy. The common Halkier/Helgaker complete basis set (CBS) extrapolation[104,105] based on the def2-TZVPP and def2-QZVPP basis sets[53] was carried out with corresponding auxiliary basis sets def2-TZVPP/C and def2-QZVPP/C.[106] For the two reactions showing the largest deviations for many of the tested density functional approximations, refined reference values at CCSD(T)/aug-cc-pwCVTZ[107] level are presented in Ref. 102. The accuracy of the reference values is conservatively estimated to ±2 kcal/mol.

Statistical results and one representative reaction are shown in Fig. 7. As expected, methods lacking semi-local electron correlation like HF-3c cannot describe thermochemistry properly (similar to the GMTKN55 thermochemistry subsets). PBEh-3c and B97-3c are both performing well with MADs of 5.2 and 4.3 kcal/mol, i.e., B97-3c is the most reliable of our "3c" methods for metal organic reactions. Indeed, it is better than both B97-D3 and the popular B3LYP-D3 functional with MADs of 6.0 and 5.0 kcal/mol (evaluated in an extended def2-QZVPP basis set but without the D3 ATM term).[102] Especially, reactions of larger reactants like the σ-donor complexation depicted in Fig. 7 are described well by B97-3c with deviation to the CCSD(T) reference below 4%. Due to the bulky substituents, the London dispersion contribution is crucial amounting to 93% of the binding energy. B97-3c shows the largest deviation (−16.0 kcal/mol) from the reference value (−14.7 kcal/mol) for the ligand exchange reaction of the [(η^6-Bz)Mo(CO)$_3$] complex with MeCN involving the cleavage of the η^6-bound benzyl moiety, but this reaction energy is drastically overestimated by almost all tested density functional approximations.[102] However, the overall performance of B97-3c for transition metal chemistry is promising. Particularly, when keeping in mind that it is one and two orders of magnitude faster than typical GGAs and hybrids with larger quadruple-ζ basis sets, respectively, B97-3c is a good candidate for large-scale studies of metal-organic reactions.

E. Lattice energies and geometries of molecular crystals

Molecular crystals are an important class of systems with an impact on material science in general[108] and on the pharmaceutical industry in particular.[32,109] While individual systems can be described by high-level methods with a high computational effort, this has been only possible for a few selected systems.[25,110,111] Efficient DFT methods are crucial for routine applications and they indeed show promising accuracies for the stability ranking of competing polymorphs.[112] Thus, it is worthwhile testing if our composite methods with reduced basis set expansions are still applicable. The minimal basis set HF-3c method was originally developed for molecular complexes only and later tests showed good lattice energies but systematically too large crystal mass densities.[113,114] Thus, a slight modification with scaled dipole-quadrupole London dispersion contribution, dubbed sHF-3c, has been proposed.[115] The s_8 scaling factor has been reduced by 30%, still yielding good interaction energies, but providing a more balanced description between energetic and geometric properties. It has been specifically tested for a broad range of molecular crystals,[115] which showed promising results in a recent study,[116] and is thus used for comparison in this subsection. Instead of PBEh-3c, we use the screened exchange variant HSE-3c as it

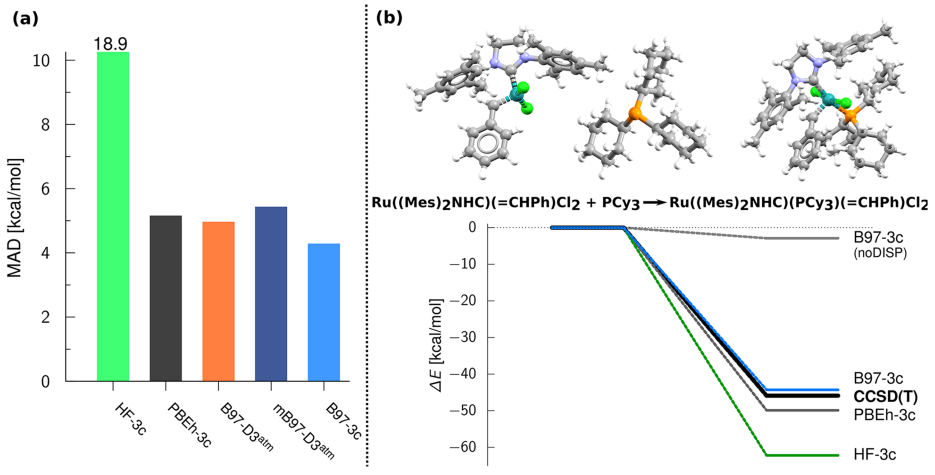

FIG. 7. (a) Mean absolute deviations of various methods for metal organic reactions of the MOR benchmark set. (b) Molecular systems and reaction energy of the largest σ-donor complexation from the MOR set (reaction 25). All B97 variants have been evaluated in the same mTZVP basis set.

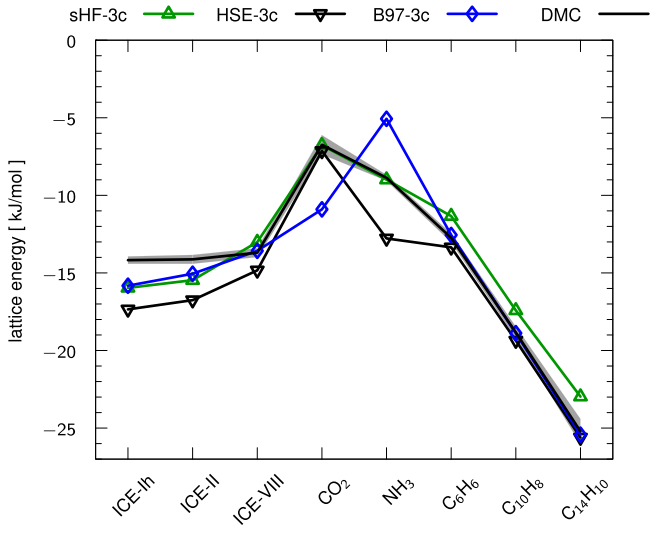

FIG. 8. Lattice energies of eight molecular crystals are compared with DMC reference energies. The gray shading shows the 2σ confidence interval of the references.

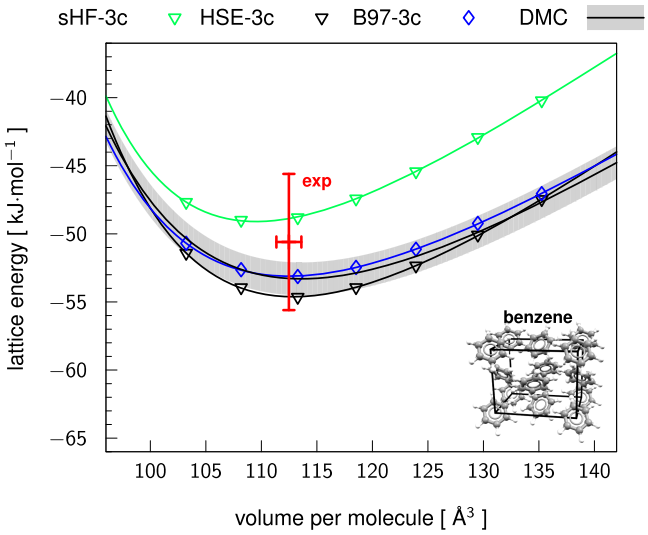

FIG. 9. Third-order Birch-Murnaghan isothermal equation of state of the benzene crystal computed by DMC and compared to composite DFT methods. The gray shading highlights the 2σ confidence interval. The inset shows the primitive unit cell, references taken from Ref. 26.

is numerically much more stable in periodic boundary conditions.[34] The central quantity to assess the stability of a crystal is its lattice energy E_{latt}, which is the energy per molecule gained upon adopting the crystalline form with respect to the gas state

$$E_{\text{latt}} = \frac{1}{Z}E_{\text{crys}} - E_{\text{gas}}, \tag{13}$$

with E_{crys} being the energy of the crystal with Z molecules in the primitive cell and E_{gas} being the energy of the isolated molecule in its lowest energy conformation. Recently, reference energies based on high-level diffusion Monte Carlo (DMC) have been calculated for a set of eight molecular crystals, namely, three ice polymorphs (Ih, II, VIII), carbon dioxide, ammonia, benzene, naphtalene, and anthracene.[26] This set is quite challenging for all electronic structure methods as it contains strong hydrogen bonds as well as unsaturated π-systems with dominant London dispersion interactions. Figure 8 shows the results of the "3c" composite schemes. We see some larger deviations for carbon dioxide and ammonia, which have been known to be challenging for GGA functionals.[7] The π-stacked systems are unproblematic for all our composite methods, and sHF-3c seems to have an excellent compensation of errors for the chosen systems as even the ice polymorphs are comparably well behaved. B97-3c is even slightly better for ice, which can be tracked back to the repulsive character of the semi-local exchange correlation (similar to revPBE or BLYP) in combination with the double polarization functions on all oxygen atoms. The MADs of sHF-3c, HSE-3c, and B97-3c are quite comparable with 1.1, 1.3, and 1.4 kcal/mol and outperform more elaborate methods like the random phase approximation (RPA@PBE) that yields an MAD of 1.9 kcal/mol.[26] In addition to the single-point lattice energy, a potential energy curve has been constructed for benzene, the comparison to the "3c" methods is shown in Fig. 9. The B97-3c potential surface and all computed properties (lattice energy, equilibrium geometry, and bulk modulus) are in very close agreement with the high-level references. One can appreciate the merit of the theoretical reference as the semi-experimental values have a quite substantial uncertainty and all shown methods would be within its error margins.

Statistics on eight systems are of course of limited use and we thus additionally analyze an extended set of 33 molecular crystals from the ICE10[117] and X23[118,119] lattice energy benchmark sets. Here, semi-experimental references are derived from measured sublimation enthalpies (ΔH_{sub}) by

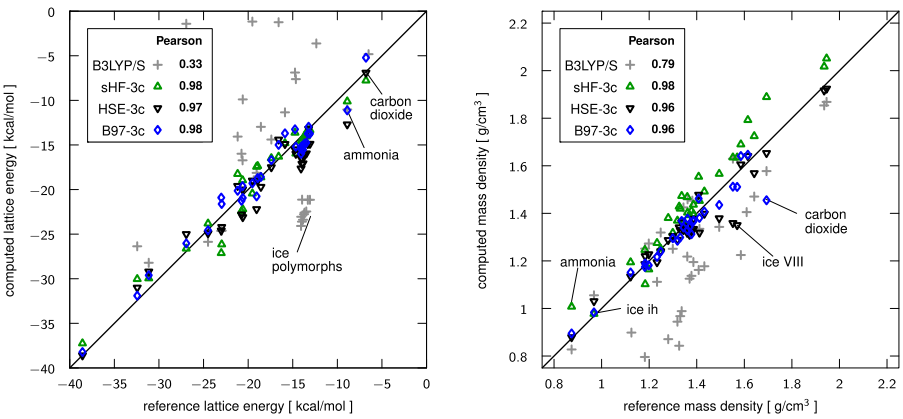

FIG. 10. Lattice energies and mass densities of 33 molecular crystals from the ICE10[117] and X23[118,119] benchmark set. In addition to the three composite methods, we show results from B3LYP in a small basis set taken from Ref. 15 and give the Pearson linear correlation value.

TABLE VII. Statistical deviations of lattice energies ΔE_{latt} and unit cell volumes ΔV_{cell} for 33 molecular crystals in the ICE10 and X23 benchmark sets. Negative MDs indicate crystals that are bound too strongly or are too dense, respectively.

	ICE10 ΔE_{latt} (kcal/mol)				ICE10 ΔV_{cell} (Å^3)			
	MD	MAD	SD	AMAX	MD	MAD	SD	AMAX
sHF-3c	−0.6	0.6	0.5	1.6	−0.1	0.6	0.8	1.9
HSE-3c	−2.6	2.6	0.6	3.6	0.4	1.1	1.5	3.1
B97-D3$^{\text{atm}}$ [a]	−1.7	1.7	0.3	2.2	0.2	0.3	0.3	0.7
B97-3c[a]	−1.2	1.2	0.4	1.9	0.1	0.3	0.4	0.7
	X23 ΔE_{latt} (kcal/mol)				**X23** ΔV_{cell} [Å^3]			
sHF-3c	0.1	1.4	1.7	4.1	−6.0	6.0	3.7	16.0
HSE-3c	−0.3	1.3	1.6	3.8	1.2	3.5	3.2	8.4
B97-D3$^{\text{atm}}$ [a,b]	−1.3	1.4	1.2	3.3	−0.1	2.3	2.9	6.1
B97-3c[a,b]	0.5	1.0	1.1	2.2	0.9	2.6	2.9	7.1

[a]Evaluated in the same mTZVP basis set.
[b]Five systems treated as single-points on HSE-3c structures due to numerical problems (see the supplementary material).

removing zero-point (E_{ZPVE}) and thermal effects

$$E_{\text{latt}} = -\Delta H_{\text{sub}}(T) + \Delta E_{\text{ZPVE}} + \int_0^T \Delta C_p(T')\,\mathrm{d}T', \qquad (14)$$

with measurement temperature T and isobaric heat capacity differences ΔC_p. It is important to keep in mind that these references have an uncertainty of about 1-1.5 kcal/mol mostly due to uncertainties in the sublimation measurement.[120] The zero-point and thermal effect on the unit cell volumes have been estimated for ICE10[117] and X23,[15] here the back-corrected volumes have an uncertainty of about 1%-2%. Individual results are shown in Fig. 10 and the statistical evaluation is summarized in Table VII. The strong hydrogen bonds in the ICE10 set are particularly challenging for small basis set expansions. Still, the scaled minimal basis set in HF-3c seems to be an ideal spot for compensating these errors. HSE-3c is a bit worse, probably due to underestimated BSSE, but still much better than other hybrid functionals in small basis set expansions as shown by B3LYP/def2-SV(P). B97-3c is performing as well as GGA functionals in converged basis sets like BLYP-D3,[117] yielding a lattice MAD of 1.2 kcal/mol and mass densities that are on average only 0.7% too small. Similar to other studies, ice Ih is the most problematic system and B97-3c overestimates its lattice energy by 1.9 kcal/mol. For the X23 set, similar trends can be seen and the lattice energy MADs of sHF-3c, HSE-3c, and B97-3c are 1.4, 1.3, and 1.0 kcal/mol, respectively. B97-3c is within the chemical accuracy, which could be previously only shown for very few dispersion corrected hybrid functionals in a large plane-wave basis set expansion.[7] The main outliers are carbon dioxide and ammonia with deviations of 1.6 and 2.2 kcal/mol, respectively, from the reference lattice energy. Excluding them from the statistics, the relative mean absolute deviations of B97-3c are 4.4% and 2.5% for the lattice energies and mass densities, respectively.

V. CONCLUSIONS

We have presented a new low-cost electronic structure method that ideally combines the semi-local B97 density functional in a medium-sized basis set expansion of triple-ζ quality with classical correction potentials for optimal performance. It is by a factor of two to three faster than a typical GGA in the standard def2-TZVP basis set and an order of magnitude faster when compared to (almost) converged def2-QZVP calculations. We have demonstrated that this gain in computational speed does not deteriorate the accuracy and B97-3c indeed even outperforms traditional GGA functional in terms of molecular geometries and noncovalent interactions for large complexes and molecular crystals. B97-3c seems to be very promising for metal organic reactions and seems to be more reliable than the widely applied B3LYP-D3 functional. Compared to our other "3c" low-cost methods, B97-3c will be most useful for systems with partial multi-reference character, where Fock exchange should be avoided. Similar to HF-3c and PBEh-3c, B97-3c can be readily applied without further correction for BSSE or London dispersion. It represents a robust high-speed electronic structure method applicable to hundreds of atoms and a single computer node. We recommend its use in a multi-level approach, i.e., for geometry optimizations and for prescreening energies.

SUPPLEMENTARY MATERIAL

See supplementary material for mTZVP basis set definition, explicit individual results of the benchmark sets CCse21, LMGB35, HMGB11, TMC32, ROT34, S66x8, S30L, L7, MOR31, GMTKN55, DMC8, ICE10, and X23.

ACKNOWLEDGMENTS

This work was supported by the DFG in the framework of the "Gottfried-Wilhelm-Leibniz" prize. J.G.B. acknowledges support by the Alexander von Humboldt foundation within the Feodor-Lynen program and A.H. acknowledges technical support by Christoph A. Bauer for evaluations of the GMTKN55 database. The authors thank Jakob Seibert for proofreading.

[1]R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989).
[2]W. Kohn, Rev. Mod. Phys. **71**, 1253 (1998).
[3]K. Burke, J. Chem. Phys. **136**, 150901 (2012).
[4]J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. **115**, 036402 (2015).
[5]N. Mardirossian and M. Head-Gordon, Phys. Chem. Chem. Phys. **16**, 9904 (2014).
[6]Y. Wang, X. Jin, H. S. Yu, D. G. Truhlar, and X. He, Proc. Natl. Acad. Sci. U. S. A. **114**, 8487 (2017).
[7]S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, Chem. Rev. **116**, 5105 (2016).
[8]J. Klimeš and A. Michaelides, J. Chem. Phys. **137**, 120901 (2012).
[9]G. J. O. Beran, Chem. Rev. **116**, 5567 (2016).
[10]Y. Zhao and D. G. Truhlar, Acc. Chem. Res. **41**, 157 (2008).
[11]R. Paverati and D. G. Truhlar, Philos. Trans. R. Soc., A **372**, 20120476 (2013).
[12]L. Goerigk and S. Grimme, J. Chem. Theory Comput. **7**, 291 (2011).
[13]N. Mardirossian, J. A. Parkhill, and M. Head-Gordon, Phys. Chem. Chem. Phys. **13**, 19325 (2011).
[14]N. Mardirossian and M. Head-Gordon, Mol. Phys. **115**, 2315 (2017).
[15]S. Grimme, G. Brandenburg, C. Bannwarth, and A. Hansen, J. Chem. Phys. **143**, 054107 (2015).
[16]J. Witte, M. Goldey, J. B. Neaton, and M. Head-Gordon, J. Chem. Theory Comput. **11**, 1481 (2015).
[17]M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko, and V. Barone, J. Phys. Chem. A **119**, 2058 (2015).

[18]W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (Wiley-VCH, New York, 2001).
[19]C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, and F. Neese, J. Chem. Phys. **144**, 024109 (2016).
[20]S. A. Maurer, D. S. Lambrecht, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **138**, 014101 (2013).
[21]M. Schütz, O. Masur, and D. Usvyat, J. Chem. Phys. **140**, 244107 (2014).
[22]A. Zen, S. Sorella, M. J. Gillan, A. Michaelides, and D. Alfè, Phys. Rev. B **93**, 241118(R) (2016).
[23]C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, J. Chem. Phys. **139**, 134101 (2013).
[24]A. Ambrosetti, D. Alfè, R. A. DiStasio, and A. Tkatchenko, J. Phys. Chem. Lett. **5**, 849 (2014).
[25]J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, and G. K.-L. Chan, Science **345**, 640 (2014).
[26]A. Zen, J. G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè, and A. Michaelides, "Fast and accurate quantum Monte Carlo for molecular crystals," Proc. Natl. Acad. Sci. U. S. A. (in press).
[27]L. A. Curtiss, P. C. Redfern, and K. Raghavachari, J. Chem. Phys. **126**, 084108 (2007).
[28]E. Paulechka and A. Kazakov, J. Phys. Chem. A **121**, 4379 (2017).
[29]R. Sure, J. Antony, and S. Grimme, J. Phys. Chem. B **118**, 3431 (2014).
[30]J. Yin, N. M. Henriksen, D. R. Slochower, M. R. Shirts, M. W. Chiu, D. L. Mobley, and M. K. Gilson, J. Comput.-Aided Mol. Des. **31**, 1 (2017).
[31]A. M. Reilly *et al.*, Acta Crystallogr., Sect. B **72**, 439 (2016).
[32]S. L. Price and S. M. Reutzel-Edens, Drug Discovery Today **21**, 912 (2016).
[33]R. Sure and S. Grimme, J. Comput. Chem. **34**, 1672 (2013).
[34]J. G. Brandenburg, E. Caldeweyher, and S. Grimme, Phys. Chem. Chem. Phys. **18**, 15519 (2016).
[35]J. Hostaš and J. Řezáč, J. Chem. Theory Comput. **13**, 3575 (2017).
[36]J. Witte, J. B. Neaton, and M. Head-Gordon, J. Chem. Phys. **146**, 234105 (2017).
[37]A. Otero-de-la Roza and G. A. DiLabio, J. Chem. Theory Comput. **13**, 3505 (2017).
[38]K. Miyamoto, T. F. Miller, and F. R. Manby, J. Chem. Theory Comput. **12**, 5811 (2016).
[39]R. Sure, J. G. Brandenburg, and S. Grimme, ChemistryOpen **5**, 94 (2016).
[40]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[41]H. Kruse and S. Grimme, J. Chem. Phys. **136**, 154101 (2012).
[42]J. Kussmann and C. Ochsenfeld, J. Chem. Phys. **138**, 134114 (2013).
[43]I. S. Ufimtsev and T. J. Martinez, J. Chem. Theory Comput. **5**, 2619 (2009).
[44]N. Luehr, I. S. Ufimtsev, and T. J. Martinez, J. Chem. Theory Comput. **7**, 949 (2011).
[45]S. Ehrlich, A. H. Göller, and S. Grimme, ChemPhysChem **18**, 898 (2017).
[46]S. Grimme, J. Comput. Chem. **27**, 1787 (2006).
[47]N. Mardirossian and M. Head-Gordon, J. Chem. Phys. **142**, 074111 (2015).
[48]A. D. Becke, J. Chem. Phys. **107**, 8554 (1997).
[49]H. L. Schmider and A. D. Becke, J. Chem. Phys. **108**, 9624 (1998).
[50]A. D. Boese, N. L. Doltsinis, N. C. Handy, and M. Sprik, J. Chem. Phys. **112**, 1670 (2000).
[51]T. W. Keal and D. J. Tozer, J. Chem. Phys. **123**, 121103 (2005).
[52]A. D. Boese and J. M. L. Martin, J. Chem. Phys. **121**, 3405 (2004).
[53]F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).
[54]B. M. Axilrod and E. Teller, J. Chem. Phys. **11**, 299 (1943).
[55]Y. Muto, Proc. Phys.-Math. Soc. Jpn. **17**, 629 (1943).
[56]R. A. DiStasio, V. V. Gobre, and A. Tkatchenko, J. Phys.: Condens. Matter **26**, 213202 (2014).
[57]J. F. Dobson, Int. J. Quantum Chem. **114**, 1157 (2014).
[58]M. R. Kennedy, A. R. McDonald, A. E. DePrince, M. S. Marshall, R. Podeszwa, and C. D. Sherrill, J. Chem. Phys. **140**, 121104 (2014).
[59]S. Grimme, S. Ehrlich, and L. Goerigk, J. Comput. Chem. **32**, 1456 (2011).
[60]E. R. Johnson and A. D. Becke, J. Chem. Phys. **123**, 024101 (2005).
[61]E. R. Johnson and A. D. Becke, J. Chem. Phys. **124**, 174104 (2006).
[62]J. G. Brandenburg, J. E. Bates, J. Sun, and J. P. Perdew, Phys. Rev. B **94**, 115144 (2016).
[63]K. Levenberg, Q. Appl. Math. **2**, 164 (1944).
[64]D. W. Marquardt, SIAM J. Appl. Math. **11**, 431 (1963).
[65]L. Goerigk, A. Hansen, C. A. Bauer, S. Ehrlich, A. Najibi, and S. Grimme, Phys. Chem. Chem. Phys. **19**, 32184 (2017).
[66]See https://www.chemie.uni-bonn.de/pctc/mulliken-center/software for the SRB correction within the gCP standalone code.
[67]F. Neese, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 73 (2012).
[68]R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, Chem. Phys. Lett. **162**, 165 (1989).
[69]TURBOMOLE V7.0 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
[70]R. Dovesi, R. Orlando, A. Erba, C. M. Zicovich-Wilson, B. Civalleri, S. Casassa, L. Maschio, M. Ferrabone, M. De La Pierre, P. D'Arco, Y. Noël, M. Causà, M. Rérat, and B. Kirtman, Int. J. Quantum Chem. **114**, 1287 (2014).
[71]The implementation was done by Dr. Eva Perlt from the University of Bonn. A slightly different basis set is employed; see E. Perlt, P. Ray, A. Hansen, F. Malberg, S. Grimme, and B. Kirchner, "Finding the best density functional approximation to describe interaction energies and structures of ionic liquids in molecular dynamics studies," J. Chem. Phys. (submitted).
[72]The CP2K Developers Group, 2015, CP2K is freely available from http://www.cp2k.org/.
[73]J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 15 (2014).
[74]E. J. Baerends, D. E. Ellis, and P. Ros, Chem. Phys. **2**, 41 (1973).
[75]B. I. Dunlap, W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).
[76]K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, Chem. Phys. Lett. **240**, 283 (1995).
[77]F. Weigend, Phys. Chem. Chem. Phys. **8**, 1057 (2006).
[78]K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, Theor. Chem. Acc. **97**, 119 (1997).
[79]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); Erratum **78**, 1396 (1997).
[80]J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
[81]A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[82]J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).
[83]E. Brémond, M. Savarese, N. Q. Su, A. J. Pérez-Jiménez, X. Xu, J. C. Sancho-García, and C. Adamo, J. Chem. Theory Comput. **12**, 459 (2016).
[84]J. A. Pople, *Modern Theoretical Chemistry* (Plenum, New York, 1976), Vol. 4.
[85]Note, however, that SCAN-D3/def2-QZVP[4,62] has an even smaller MAD for the CCse21 bond lengths of 0.37 pm and yields a small average structure RMSD of 0.42 pm and is thus the most accurate standard (non-double-hybrid) functional for geometries of small semi-rigid molecules.
[86]M. Bühl and H. Kabrede, J. Chem. Theory Comput. **2**, 1282 (2006).
[87]S. Grimme, J. Chem. Phys. **124**, 034108 (2006).
[88]S. Grimme and M. Steinmetz, Phys. Chem. Chem. Phys. **15**, 16031 (2013).
[89]T. Risthaus, M. Steinmetz, and S. Grimme, J. Comput. Chem. **35**, 1509 (2014).
[90]On the other hand, both M06L and SCAN converge slow with respect to the single particle basis set and require large radial grid sizes.[62] In practice, B97-3c is about 2 orders of magnitude faster as confirmed for a standard single-point energy of phenanthrene and [7]helicene. B97-3c is used in default settings (see Sec. III B), while M06L is evaluated with RI-J in a def2-QZVP basis set and increased gridsize (*grid7*, IntAcc = 33.5), i.e., with ≥400 radial grid points per atom in following the recommendation in Ref. 62. With these settings, a single-point calculation with M06L of phenanthrene and [7]helicene takes 01:45:46 and 09:42:35, respectively (hh:mm:ss).
[91]K. Pernal, R. Podeszwa, K. Patkowski, and K. Szalewicz, Phys. Rev. Lett. **103**, 263201 (2009).
[92]É. D. Murray, K. Lee, and D. C. Langreth, J. Chem. Theory Comput. **5**, 2754 (2009).
[93]J. Řezáč, K. E. Riley, and P. Hobza, J. Chem. Theory Comput. **7**, 2427 (2011).
[94]B. Brauer, M. K. Kesharwani, S. Kozuch, and J. M. L. Martin, J. Chem. Theory Comput. **18**, 20905 (2016).
[95]R. Sure and S. Grimme, J. Chem. Theory Comput. **11**, 3785 (2015).
[96]R. Sedlak, T. Janowski, M. Pitoňák, J. Řezáč, P. Pulay, and P. Hobza, J. Chem. Theory Comput. **9**, 3364 (2013).
[97]T. Risthaus and S. Grimme, J. Chem. Theory Comput. **9**, 1580 (2013).
[98]H. Kruse, A. Mladek, K. Gkionis, A. Hansen, S. Grimme, and J. Sponer, J. Chem. Theory Comput. **11**, 4972 (2015).
[99]Y. A. Aoto, A. P. de Lima Batista, A. Köhn, and A. G. S. de Oliveira-Filho, J. Chem. Theory Comput. **13**, 5291 (2017).

[100]J. J. Determan, K. Poole, G. Scalmani, M. J. Frisch, B. G. Janesko, and A. K. Wilson, J. Chem. Theory Comput. **13**, 4907 (2017).
[101]C. J. Cramer and D. G. Truhlar, Phys. Chem. Chem. Phys. **11**, 10757 (2009).
[102]S. Dohm, M. Steinmetz, M. P. Checinski, A. Hansen, and S. Grimme, "Comprehensive thermochemical benchmark set of realistic metal organic reactions," J. Chem. Theory Comput. (submitted).
[103]F. Pavošević, C. Peng, P. Pinski, C. Riplinger, F. Neese, and E. F. Valeev, J. Chem. Phys. **146**, 174108 (2017).
[104]A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, and J. Olsen, Chem. Phys. Lett. **302**, 437 (1999).
[105]A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).
[106]A. Hellweg, C. Hättig, S. Höfener, and W. Klopper, Theor. Chem. Acc. **117**, 587 (2007).
[107]B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, Jr., and A. K. Wilson, Theor. Chem. Acc. **128**, 69 (2011).
[108]A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. A. Little, S. Y. Chong, B. J. Slater, D. P. McMahon, B. Bonillo, C. J. Stackhouse, A. Stephenson, C. M. Kane, R. Clowes, T. Hasell, A. I. Cooper, and G. M. Day, Nature **543**, 657 (2017).
[109]A. J. Cruz-Cabeza, S. M. Reutzel-Edens, and J. Bernstein, Chem. Soc. Rev. **44**, 8619 (2015).
[110]M. J. Gillan, D. Alfè, P. J. Bygrave, C. R. Taylor, and F. R. Manby, J. Chem. Phys. **139**, 114101 (2013).
[111]S. Wen and G. J. O. Beran, J. Chem. Theory Comput. **7**, 3733 (2011).
[112]J. G. Brandenburg and S. Grimme, Acta Crystallogr., Sect. B **72**, 502 (2016).
[113]J. G. Brandenburg and S. Grimme, Top. Curr. Chem. **345**, 1 (2014).
[114]J. G. Brandenburg, J. Potticary, H. A. Sparkes, S. L. Price, and S. R. Hall, J. Phys. Chem. Lett. **8**, 4319 (2017).
[115]M. Cutini, B. Civalleri, M. Corno, R. Orlando, J. G. Brandenburg, L. Maschio, and P. Ugliengoa, J. Chem. Theory Comput. **12**, 3340 (2016).
[116]M. Cutini, M. Corno, and P. Ugliengo, J. Chem. Theory Comput. **13**, 370 (2017).
[117]J. G. Brandenburg, T. Maas, and S. Grimme, J. Chem. Phys. **142**, 124104 (2015).
[118]A. Otero-de-la-Roza and E. R. Johnson, J. Chem. Phys. **137**, 054103 (2012).
[119]A. M. Reilly and A. Tkatchenko, J. Chem. Phys. **139**, 024705 (2013).
[120]J. S. Chickos, Netsu Sokutei **30**, 116 (2003).